

Scaling detection in time series: diffusion entropy analysis

Nicola Scafetta¹, and Paolo Grigolini^{1,2,3}.

¹Center for Nonlinear Science, University of North Texas, P.O. Box 311427, Denton, Texas 76203-1427

²Dipartimento di Fisica dell'Università di Pisa and INFN, Piazza Torricelli 2, 56127 Pisa, Italy

³Istituto di Biofisica CNR, Area della Ricerca di Pisa, Via Alfieri 1, San Cataldo 56010 Ghezzano-Pisa, Italy

(February 1, 2008)

The methods currently used to determine the scaling exponent of a complex dynamic process described by a time series are based on the numerical evaluation of variance. This means that all of them can be safely applied only to the case where ordinary statistical properties hold true even if strange kinetics are involved. We illustrate a method of statistical analysis based on the Shannon entropy of the diffusion process generated by the time series, called Diffusion Entropy Analysis (DEA). We adopt artificial Gauss and Lévy time series, as prototypes of ordinary and anomalous statistics, respectively, and we analyse them with the DEA and four ordinary methods of analysis, some of which are very popular. We show that the DEA determines the correct scaling exponent even when the statistical properties, as well as the dynamic properties, are anomalous. The other four methods produce correct results in the Gauss case but fail to detect the correct scaling in the case of Lévy statistics.

05.45.T, 89.75.D, 05.40.F, 05.40

I. INTRODUCTION

Scale invariance has been found to hold empirically for a number of complex systems and the correct evaluation of the scaling exponents is of fundamental importance to assess if universality classes exist [1]. The mathematical definition of scaling is as follows. The function $\Phi(r_1, r_2, \dots)$ is termed scaling invariant, if it fulfills the property:

$$\Phi(r_1, r_2, \dots) = \gamma^a \Phi(\gamma^b r_1, \gamma^c r_2, \dots). \quad (1)$$

Eq. (1) means that if we scale all coordinates $\{r\}$ by means of an appropriate choice of the exponents a, b, c, \dots , then we always recover the same scaling function. The theoretical and experimental search for the correct scaling exponents is intimately related to the discovery of deviations from ordinary statistical mechanics. This aspect emerges clearly, for instance, from Ref. [2]. The author of this interesting book, with the help of dimensional analysis and regularity assumption, determines the values of the scaling exponents. These scaling exponents, however, are *trivial* in the sense that they refer to ordinary statistical mechanics.

In this paper we focus on the scaling of time series, and consequently [3] on the scaling of diffusion processes. In fact, according to the prescription of Ref. [3] we interpret the numbers of a time series as generating diffusion fluctuations, thereby shifting our attention from the time series to the probability distribution function (pdf) $p(x, t)$, where x denotes the variable collecting the fluctuations. In this case, if the time series is stationary, the scaling property of Eq. (1) takes the form

$$p(x, t) = \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right), \quad (2)$$

where δ is the scaling exponent. The ordinary value [2] of the scaling exponent is $\delta = 1/2$. Ordinary statistical mechanics is intimately related to the Central Limit Theorem (CLT) [4], thereby implying for the function F of Eq. (2) the Gaussian form.

The main purpose of this paper is to prove that a technique of statistical analysis, recently introduced to establish the thermodynamic nature of a time series of sociological interest [5], affords a reliable way to evaluate the scaling exponent. This method of analysis is based on the entropy of the diffusion process and for this reason is called Diffusion Entropy Analysis (DEA). We compare the DEA to the Standard Deviation Analysis (SDA) [6], to the Detrended Fluctuation Analysis (DFA) [3], to the Rescaled Range Analysis (RRA) [7], to the Spectral Wavelet Analysis (SWA) [8], and we show that, while all these techniques, some of which are very popular, can yield wrong scaling exponents, the DEA always determines the correct value, with satisfactory precision. This important conclusion is reached by examining artificial sequences generating both Gauss and Lévy statistics. The DEA is the only technique yielding the correct scaling in both cases. The other techniques produce correct results only in the Gauss case but fail to detect the correct scaling in the case of Lévy statistics.

II. DEA

It is remarkably simple to determine the scaling parameter δ using DEA. First of all, we transform the time series into a diffusion process with a given pdf $p(x, t)$ (Section 4 illustrates an algorithm to do that). Then, we measure the Shannon entropy

$$S(t) = - \int_{-\infty}^{+\infty} dx p(x, t) \ln[p(x, t)]. \quad (3)$$

Let us suppose that $p(x, t)$ fits the scaling condition of Eq.(2) and let us plug Eq.(2) into Eq.(3). After an easy algebra, based on changing the integration variable from x into $y = x/t^\delta$, we obtain

$$S(\tau) = A + \delta \tau, \quad (4)$$

where

$$A \equiv - \int_{-\infty}^{\infty} dy F(y) \ln[F(y)], \quad (5)$$

and

$$\tau \equiv \ln(t). \quad (6)$$

It is evident that the origin of this logarithmic time, $\tau = 0$, is related to $t = 1$, and this, in turn, depends on the time unit adopted. However, this arbitrary choice in no way affects the scaling parameter, which is given by the slope of the straight line $S(\tau)$ of Eq. (3). The adoption of a different time unit would change a given straight line into a new one, parallel to the original, and thus bearing the same scaling parameter δ .

III. GAUSS AND LÉVY DIFFUSION

The rigorous definition of diffusion scaling is that of Eq.(2). Frequently, the scaling property is also expressed by means of

$$x \propto t^\delta. \quad (7)$$

Let us see now why the scaling parameter δ is often evaluated through the second moment of the pdf $p(x, t)$. In the long-time limit the variable $x(t)$ collecting the fluctuations $\xi(t)$ has a time evolution equivalent to

$$\dot{x}(t) = \xi(t), \quad (8)$$

By time integration we get

$$x(t) = x(0) + \int_0^t dt' \xi(t'). \quad (9)$$

Let us imagine a set of infinitely many trajectories of the type of that of Eq. (9). As to the second moment $\langle x^2(t) \rangle$, we evaluate its time evolution by squaring Eq.(9) and averaging over all the trajectories of this set. Under the assumption that the process is stationary and that $\langle \xi(t) \rangle = 0$, it is straightforward to obtain

$$\langle x^2(t) \rangle = \langle x^2(0) \rangle + 2 \langle \xi^2 \rangle \int_0^t dt_1 \int_0^{t_1} dt_2 \Phi_\xi(|t_2|). \quad (10)$$

Note that to get this result we use the equilibrium correlation function

$$\Phi_\xi(t_1, t_2) \equiv \frac{\langle \xi(t_1) \xi(t_2) \rangle}{\langle \xi^2 \rangle}. \quad (11)$$

Under the stationary condition, this correlation function depends only on the time difference, namely, $\Phi_\xi(t_1, t_2) = \Phi_\xi(|t_1 - t_2|)$, and this property, with the help of a suitable change of integration variables, yields Eq.(10).

What is the connection between second moment and scaling? Having in mind Eq.(7) one would be tempted to make the conjecture that

$$\langle x^2(t) \rangle \propto t^{2\delta}. \quad (12)$$

However, this conjecture is not correct in general, and to be more rigorous let us replace Eq.(12) with

$$\langle x^2(t) \rangle \propto t^{2H}. \quad (13)$$

The adoption of the symbol H rather than the symbol δ is dictated by two main reasons. First of all, H is the symbol frequently used in literature to denote the scaling parameter. In fact, the pioneer work of Mandelbrot, in the special case of Fractional Brownian Motion (FBM) [9], identified the Hurst coefficient H [7] with the scaling parameter. This is certainly correct, but the validity of this conclusion is confined to the Gaussian case. In general, as we shall see, the asymptotic behavior of Eq.(13) does not imply that $\delta = H$. This is the second reason why we prefer to use Eq.(13) rather than Eq.(12). In conclusion, we denote by δ a parameter implying the property of Eq.(2) and by H the value corresponding to the property of Eq.(13). We shall see that even when $H \neq \delta$, the parameter H might have a physical meaning of some interest, even if it is not the real scaling.

To prove under which conditions the equality $\delta = H$ applies, and consequently Eq. (12), as well as Eq.(13), is correct, let us notice that under the assumption that the fluctuation $\xi(t)$ is Gaussian, and with no other assumption, we can prove that the pdf $p(x, t)$ fulfills the following diffusion equation

$$\frac{\partial p(x, t)}{\partial t} = D(t) \frac{\partial^2}{\partial x^2} p(x, t), \quad (14)$$

where

$$D(t) \equiv \langle \xi^2 \rangle \int_0^t \Phi_\xi(t') dt'. \quad (15)$$

The proof of this important result rests on the cumulant theory of Ref. [10] and the interested reader can derive it from a more general case discussed in Ref. [11]. It is straightforward to show that the general solution of Eq. (14), for a set of particles initially located at $x = 0$, is

$$p(x, t) = \frac{1}{\sqrt{2\pi \langle x^2(t) \rangle}} \exp\left(-\frac{x^2}{2 \langle x^2(t) \rangle}\right), \quad (16)$$

where $\langle x^2(t) \rangle$ is the second moment with the time evolution described by Eq.(10). It is easy to show that the time asymptotic properties of the second moment are compatible with Eq.(13), with H ranging from 0 to 1, in the case where

$$\lim_{t \rightarrow \infty} \Phi_\xi(t) = \text{sign} \frac{\text{const}}{t^\beta}, \quad (17)$$

with

$$H = 1 - \frac{\beta}{2} \quad (18)$$

and $0 \leq \beta \leq 1$, if $\text{sign} = 1$, and $1 \leq \beta \leq 2$, if $\text{sign} = -1$. It is evident that in this physical condition the property of Eq.(2) applies to the time asymptotic limit, and consequently in the same time limit the property $\delta = H$ holds true. In Section 3A we shall shortly review the FBM, which is a mathematical idealization fitting the condition $\delta = H$. In Section 3B we shall illustrate the Lévy flight diffusion, which is, on the contrary, a mathematical idealization incompatible with the existence of a finite second moment, thereby implying a striking violation of $\delta = H$. Finally, in Section 3C we illustrate the Lévy walk diffusion, a compromise between the two earlier conditions, in the sense that it is compatible with a finite second moment. However, as we shall see, also the Lévy walk diffusion results in a violation of $\delta = H$.

A. Fractional Brownian Motion

Fractional Brownian Motion, FBM, is a generalization of ordinary Brownian motion introduced by Mandelbrot [9] to describe anomalous diffusion. By definition, the FBM of index η is described by the fractional Gaussian propagator

$$p(x, t) = \frac{1}{\sqrt{4\pi D t^\eta}} \exp\left(-\frac{(x - \bar{x})^2}{4D t^\eta}\right). \quad (19)$$

The variance of this pdf is given by

$$\langle (x - \bar{x})^2 \rangle = \int_{-\infty}^{\infty} dx (x - \bar{x})^2 p(x, t) = 2D t^\eta. \quad (20)$$

We see that FBM is compatible with the asymptotic limit of the dynamic approach to diffusion generated by Eq. (8) when the fluctuation $\xi(t)$ is Gaussian. Furthermore, we see that Eq.(19) fits the scaling condition of Eq. (2) and Eq.(20) fits the second moment scaling of Eq. (13). Thus we conclude that

$$\delta = \frac{\eta}{2} = H. \quad (21)$$

In other words, FBM is a kind of diffusion where the second moment scaling mirrors correctly the scaling property of Eq.(2).

B. Lévy Flight Diffusion

Lévy Flight Diffusion is described, in the symmetric case, by the characteristic function:

$$\hat{p}(k, t) = \exp(-K^\alpha t |k|^\alpha). \quad (22)$$

This type of diffusion is generated by a walker that makes jumps with lengths determined by a probability density function, $\lambda(\xi)$, whose asymptotic behavior function has the following inverse power law form:

$$\lambda(\xi) \sim A_\alpha \sigma^\alpha |\xi|^{-1-\alpha} = A_\alpha \sigma^{1-\mu} |\xi|^{-\mu}. \quad (23)$$

This means $|\xi| \gg \sigma$ and $\mu = 1 + \alpha$.

The Fourier inversion of (22) can be obtained analytically by making use of the Fox function [12,13]

$$p(x, t) = \frac{1}{\mu - 1} \frac{1}{t^{1/(\mu-1)}} \left(\frac{|x|}{t^{1/(\mu-1)}} \right)^{-1} H_{2,2}^{1,1} \left[\frac{|x|}{(K^\alpha t)^{1/(\mu-1)}} \mid \begin{matrix} (1, 1/\alpha), (1, 1/2) \\ (1, 1), (1, 1/2) \end{matrix} \right].$$

It is evident that this expression fits the scaling definition of Eq. (2), and that consequently, for $1 < \mu < 3$, the scaling coefficient δ is

$$\delta = \frac{1}{\mu - 1}. \quad (24)$$

It is important to remark that while the CLT yields Eq.(2) with $\delta = \frac{1}{2}$ with F being a Gaussian function, the Generalized Central Limit Theorem (GCLT) [14] yields Eq.(2) with $\delta \neq \frac{1}{2}$ and F departing from the Gaussian form. This departure from Gaussian statistics has the striking consequence of making the second moment diverge. In fact, in the asymptotic limit of large x 's, we get for the distribution $p(x, t)$ the following inverse power law form

$$\lim_{|x| \rightarrow \infty} p(x, t) \approx \frac{1}{t^{1/(\mu-1)}} \left(\frac{|x|}{t^{1/(\mu-1)}} \right)^{-\mu} \quad \mu < 3. \quad (25)$$

This makes the second moment $\langle x^2 \rangle$ diverge. In the more general asymmetric case [14] similar inverse power law properties apply to one of the two tails, thereby making the variance $\langle (x - \bar{x})^2 \rangle$ diverge. Consequently, the variance method for scaling detection is quite inappropriate in this case. In the case of real data, the experimental observation is done with a finite number of data. This means that the variance is finite, thereby giving the impression that the variance method can be safely adopted as a scaling detector. This would lead to misleading results, determined by the lack of accurate statistics rather than by the genuine statistical properties of the system under study.

C. A compromise: Lévy Walk Diffusion

Let us now consider another random walk prescription, Lévy Walk Diffusion [15]. This has to do with drawing the random numbers τ_i 's, with the probability density $\psi(\tau)$ given by

$$\psi(\tau) = (\mu - 1) \frac{T^{\mu-1}}{(T + \tau)^\mu}. \quad (26)$$

This means that we can build up an infinite sequence of numbers, which are then used to make a random walker walk with the following rules. At time $t = 0$, when the first random number, τ_1 , is selected, we also toss a coin to decide whether the random walker has to move in the positive or in the negative direction. The random walker walks with a velocity of constant intensity W . Thus, tossing a coin serves the purpose of establishing whether the velocity of the random walker is W (head) or $-W$ (tail). This condition of uniform motion lasts for a time interval of duration τ_1 . At the end of this condition of uniform motion, a new number, τ_2 , is randomly drawn, and a new velocity direction is established by another coin tossing. The time series $\{\tau_i\}$ is converted into a time series $\{\xi_i\}$ of either 1's or -1 's, as follows. We associate each τ_i to a patch of n_i identical symbols, the i -th patch. The values of these symbols are fixed to be either 1's or -1 's, according to the coin tossing rule. Then we sew the i -th patch to the $i - 1$ -th, on the left, and to the $i + 1$ -th, on the right, thereby creating a virtually infinite sequence of ξ_i , with values given by either 1 or -1 . It is evident that the variable x collecting the fluctuations ξ_i corresponds to the earlier walking prescription with $W = 1$. The error resulting from the adoption of only the integer part of τ_i is made irrelevant by the fact that the inverse power law nature of Eq. (26) with $\mu < 3$ makes statistically significant the numbers $\tau_i \gg 1$. It is worthe remarking that we shall focus our attention on the condition $2 < \mu < 3$.

It is important to stress that the physical condition $\mu > 2$ corresponds to the non-vanishing mean time $< \tau >$, whose explicit expression is:

$$< \tau > = \frac{T}{\mu - 2}. \quad (27)$$

We denote as *event* the joint process of random drawing of a number and of coin tossing. We consider a time scale characterized by the property

$$t \gg \tau. \quad (28)$$

It is evident that the number of events that occurred prior to a given time t is given by

$$n = \frac{t}{< \tau >}. \quad (29)$$

Consequently, at a given time $t \gg < \tau >$, the position x occupied by the random walker is equivalent to the superposition of many highly correlated fluctuations ξ_i or of n uncorrelated variables η_i , whose modulus is equal to $W\tau_i$, with signs randomly assigned by the coin tossing. Thus, the probability distribution function, $\lambda(\eta)$, is given by

$$\lambda(\eta) = \frac{1}{2W} \psi\left(\frac{\eta}{W}\right), \quad (30)$$

the analytical form of the function ψ being given by Eq.(26). The variables η_i can be identified with the variables ξ of Section 3B. As a consequence, the asymptotic properties of this probability distribution function are the same as those of Eq.(23). By applying again the GCLT [14] we obtain the same diffusion process as that of Eq.(22), and, of course, the same scaling prescription as that of Eq.(24).

This walk prescription was termed Symmetric Velocity Model in Ref. [15] and is a form of Lévy walk. The Lévy walk has been originally introduced [16–18] as a dynamic approach to Lévy statistics more realistic than the Lévy flight of Section 3B. The reason for this conviction is that with the Lévy walk the walker makes a step of a given length in a time proportional to this length, while with the Lévy flight the walker makes instantaneous jumps of arbitrarily length, a property judged, in fact, to be somewhat unrealistic. In this paper, the Lévy walk serves the very useful purpose of explaining why the emergence of Lévy statistics does not imply a total failure of the methods relating scaling to variance. In this case, in fact, the second moment is finite, and this property does not depend on the lack of sufficient statistics. It depends on the fact that no jump can occur with a length of intensity larger than Wt . In this specific case the renewal theory [19] prescribes that the correlation function $\Phi_\xi(t)$ is related to the waiting time distribution $\psi(\tau)$ by the important equation

$$\Phi_\xi(t) = \frac{1}{< \tau >} \int_0^{+\infty} (t' - t) \psi(t') dt'. \quad (31)$$

From this important relation, using Eq.(26), we derive the following analytical expression for $\Phi_\xi(t)$:

$$\Phi_\xi(t) = \left(\frac{T}{t + T} \right)^\beta. \quad (32)$$

with

$$\beta = \mu - 2. \quad (33)$$

At this stage we are equipped to derive the asymptotic properties of the pdf second moment. The existence of the correlation function of Eq.(32) allows us to use again Eq. (10) so as to reach quickly the conclusion that

$$H = \frac{4 - \mu}{2}. \quad (34)$$

This result is, in fact, obtained by plugging Eq.(33) into Eq.(18). There is no reason to identify H with δ , in this case. Rather, if we trust the GCLT and, consequently, the scaling prescription of Eq.(24), we see that δ is related to H by

$$\delta = \frac{1}{3 - 2H}. \quad (35)$$

We shall prove that the DE method detects this correct scaling, whereas, of course, the methods resting on variance cannot, even if the exponent H they detect has an interesting physical meaning.

It is worth stressing that the physical meaning of H , determined by the variance method, changes from case to case and depends on both the statistics of the process and the walking rule adopted to change the time series into a diffusion process. To illustrate this issue, let us multiply the numbers τ_i by either 1 or -1 by tossing a coin. Then, let us denote the resulting numbers by the symbol ξ . We obtain a sequence indistinguishable from that discussed in Section 3B. In this case, as we shall see, in Section 6B, the resulting value of $H = 0.5$ reflects the fact that the numbers τ_i are uncorrelated and, by no means, that the resulting statistics is Gaussian.

IV. THE DIFFUSION ALGORITHM

Let us consider a sequence of M numbers

$$\xi_i, \quad i = 1, \dots, M. \quad (36)$$

The purpose of the DEA algorithm is to establish the possible existence of a scaling, either normal or anomalous, in the most efficient way as possible without altering the data with any form of detrending. First of all, let us select an integer number l , fitting the condition $1 \leq l \leq M$. This integer number will be referred to by us as “time”. For any given time l we can find $M - l + 1$ sub-sequences defined by

$$\xi_i^{(s)} \equiv \xi_{i+s}, \quad \text{with } s = 0, \dots, M - l. \quad (37)$$

For any of these sub-sequences we build up a diffusion trajectory, s , defined by the position

$$x^{(s)}(l) = \sum_{i=1}^l \xi_i^{(s)} = \sum_{i=1}^l \xi_{i+s}. \quad (38)$$

Let us imagine this position as that of a sort of Brownian particle that at regular intervals of time has been jumping forward or backward according to the prescription of the corresponding sub-sequence of Eq.(37). This means that the particle before reaching the position that

it holds at time l has been making l jumps. The jump made at the i -th step has the intensity $|\xi_i^{(s)}|$ and is forward or backward according to whether the number $\xi_i^{(s)}$ is positive or negative. We adopt the paradigm of Brownian motion for the tutorial purpose of illustrating how the diffusion algorithm work. Actually, the ultimate task of this algorithm is to express in a quantitative way the departure of the observed process from the statistical properties of ordinary Brownian motion.

We are now ready to evaluate the entropy of this diffusion process. To do that we have to partition the x -axis into cells of size $\epsilon(l)$. When this partition is made, we have to label the cells. We count how many particles are found in the same cell at a given time l . We denote this number by $N_i(l)$. Then we use this number to determine the probability that a particle can be found in the i -th cell at time l , $p_i(l)$, by means of

$$p_i(l) \equiv \frac{N_i(l)}{(M - l + 1)}. \quad (39)$$

At this stage the entropy of the diffusion process at the time l is determined and reads

$$S_d(l) = - \sum_i p_i(l) \ln[p_i(l)]. \quad (40)$$

The easiest way to proceed with the choice of the cell size, $\epsilon(l)$, is to assume it to be a fraction of the square root of the variance of the fluctuation $\xi(i)$, and consequently independent of l .

V. THE METHODS OF ANALYSIS BASED ON VARIANCE

In this section we review four methods of time series analysis. The last three methods are very popular, and are all related, to a somewhat different extent, to the first, based on the direct evaluation of variance.

A. SDA

The direct evaluation of variance, by means of SDA, is probably the most natural method of variance detection. This method was used, for instance, in Ref. [6]. The starting point is given by the diffusion algorithm of Section 4, Eq.(38). All trajectories start from the origin $x(l = 0) = 0$. With increasing time l , the sub-sequences generate a diffusion process. At each time l , it is possible to calculate the Standard Deviation (SD) of the position of the $N - l$ trajectory with the well known expression:

$$SD(l) = \sqrt{\frac{\sum_{s=1}^{N-l} (x^{(s)}(l) - \bar{x}(l))^2}{N - l - 1}}, \quad (41)$$

where $\bar{x}(l)$ is the average of the positions of the $N - l$ sub-trajectories at time l . We note that the prescription illustrated in Section 4 to define the trajectories of this diffusion process is based on overlapping windows. This means that the trajectories are not totally independent of one another. In principle, we can also adopt a non-overlapping window method. In this case the largest trajectory that we can build up with the whole sequence is divided in $M = \text{int}(N/l)$ smaller trajectories of size l (with $\text{int}(x)$ we denote the integer part of x). Thus, the quantity $SD(l)$ can be referred to many trajectories totally independent the one from the other. The advantage of using many independent trajectories is balanced, though, by the disadvantage of statistics poorer than those obtained by using the overlapping window method. Therefore, in this paper we use the method of overlapping windows.

According to the traditional wisdom of the methods based on variance, the existence of scaling is assessed by observing, with numerical methods, the following properties

$$SD(l) \propto l^H. \quad (42)$$

The exponent H is interpreted as the scaling exponent. As discussed in Section 3, there is no guarantee that this exponent coincides with the genuine scaling δ . This is the reason why with all the methods of analysis of this section we shall use the symbol H to denote the result of the statistical analysis. To assess whether this is the true scaling or not, it is necessary to use also the DEA.

B. RRA

The RRA was introduced by Hurst in 1965, in the work, *Long-Term Storage: An Experimental Study* [7], for the main purpose of studying the water storage of the Nile river. The problem was to design a reservoir, which never overflows or empties, based upon the given record of observed discharges from a lake. Let us suppose that ξ_i is the amount of water flowing from a lake to a reservoir for each year. The problem is to determine the needed capacity of the reservoir under the condition that each year the reservoir releases a volume of water equal to the mean influx. In τ years, the average influx is

$$\langle \xi \rangle_\tau = \frac{1}{\tau} \sum_{i=1}^{\tau} \xi_i. \quad (43)$$

The amount of water accumulated in the reservoir in t years is

$$x(t, \tau) = \sum_{i=1}^t (\xi_i - \langle \xi \rangle_\tau). \quad (44)$$

The reservoir neither overflows nor empties during the period of τ years if its storage capacity is larger than the difference, $R(\tau)$, between the maximum and minimum amounts of water contained in the reservoir. $R(\tau)$ is

$$R(\tau) = \max_{1 \leq t \leq \tau} x(t, \tau) - \min_{1 \leq t \leq \tau} x(t, \tau). \quad (45)$$

For getting a dimensionless value, Hurst divided $R(\tau)$ by the standard deviation $S(\tau)$ of the data during the τ years:

$$S(\tau) = \sqrt{\frac{1}{\tau} \sum_{i=1}^{\tau} (\xi_i - \langle \xi \rangle_\tau)^2}. \quad (46)$$

Hurst observed that many phenomena are very well described by the following scaling relation:

$$\frac{R(\tau)}{S(\tau)} \propto \tau^H. \quad (47)$$

The exponent H (denoted by the letter K by Hurst) was called Hurst exponent, and consequently denoted by the letter H , by Mandelbrot [9]. The work of Mandelbrot made the RRA become popular as a technique of scaling detection. In the case of the Nile river, Hurst measured an exponent $H = 0.9$. This means that the Nile is characterized by a long range persistence that requires unusually high barriers, such as the Aswân High Dam, to contain damage and rein in the floods. This paper proves that this perspective, correct in the Gaussian case, is in general misleading.

C. DFA

The DFA was introduced in 1994 by the authors of Ref. [3]. Since 1994, hundreds of papers, which analyze fractal properties of time series with the DFA, have been published. In summary DFA works as follows. Given a time sequence $\{\xi_i\}$ ($i = 1, \dots, N$), the DFA is based upon the following steps. First, the entire sequence of length N is integrated, thereby leading to

$$x(t) = \sum_{i=1}^t (\xi_i - \langle \xi \rangle), \quad (48)$$

where

$$\langle \xi \rangle = \frac{1}{N} \sum_{i=1}^N \xi_i. \quad (49)$$

Second, the time series is divided into $\text{int}(N/l)$ non-overlapping boxes. The number l , which indicates the size of the box, is an integer smaller than N . A local trend is defined for each box by fitting the data in the box. The linear least-squares fit may be done with a

polynomial function of order $n \geq 0$ [20]. Let $x_l(t)$ be the local trend built with boxes of size l . Third, a detrended walk is defined as the difference between the original walk and the local trend given by the linear least-squares fit according to the following relation

$$X_l(t) = x(t) - x_l(t) . \quad (50)$$

Finally, the mean squared displacement of the detrended walk is calculated, that is,

$$F_D^2(l) = \frac{1}{N} \sum_{t=1}^N [X_l(t)]^2 . \quad (51)$$

The application of this method of analysis to a number of complex systems (see, for instance, Refs. [3,20]) shows that

$$F_D(n) \propto l^H . \quad (52)$$

Again, according to the traditional wisdom of the methods based on variance, the exponent H is considered to be a scaling exponent. Thus the extent of the departure from the ordinary condition of Brownian motion is given by $|H - 0.5| > 0$.

D. SWA

The SWA is a new and powerful method for studying the fractal properties of variance [8]. SWA decomposes the sample variance of a time series on a *scale-by-scale* basis. Wavelet Transform makes use of scaling wavelets that have the characteristics of being localized in space and in frequencies. The wavelets are characterized by the fact that they can be localized in the space and depend upon a scaling coefficient that gives the width of the wavelet. Two typical wavelets widely used are the Haar wavelet and the Mexican hat wavelet [8]. A width coefficient τ defines the scale analyzed by the wavelet. Given a signal $\xi(u)$, the Continuous Wavelet Transform is defined by

$$W(\tau, t) = \int_{-\infty}^{\infty} \tilde{\psi}_{\tau,t}(u) \xi(u) du . \quad (53)$$

The original signal can be recovered from its Continuous Wavelet Transform via

$$\xi(u) = \frac{1}{C_{\tilde{\psi}}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W(\tau, t) \tilde{\psi}_{\tau,t}(u) dt \right] \frac{d\tau}{\tau^2} . \quad (54)$$

Finally, it is possible to prove that

$$\int_{-\infty}^{\infty} \xi^2(u) du = \frac{1}{C_{\tilde{\psi}}} \int_0^{\infty} \left[\int_{-\infty}^{\infty} W^2(\tau, t) dt \right] \frac{d\tau}{\tau^2} \equiv \int_0^{\infty} S_W(\tau) d\tau . \quad (55)$$

The function $W^2(\tau, t)/\tau^2$ defines an energy density function that decomposes the energy across different scales and times. Eq. (55) is the wavelet equivalent to the Fourier Parseval's theorem. The function $S_W(\tau)$, defined by Eq.(55), is the wavelet spectral density function that gives the contribution to the energy at the scale τ .

From Ref. [8], we derive that SWA applied to studying a noisy sequence $\{\xi_i\}$, at the scale τ , yields

$$S_W(\tau) \propto \tau^\alpha . \quad (56)$$

The exponent α is related to the variance scaling exponent H in the same way as in the conventional Fourier Analysis. Therefore, $\alpha = 2H - 1$ for the SWA of the noise, and $\alpha = 2H$ for the SWA of the integral of the noise. The connection with the methods of scaling detection based on variance is evident.

VI. ARTIFICIAL SEQUENCE ANALYSIS

In this section we verify the theoretical predictions of the previous sections about the pdf scaling exponent δ and the variance scaling exponent, H , by using artificial sequences of five million data. With the help of artificial time series, we compare the methods of analysis based on variance with the DEA. We prove that the DEA always determines the true scaling δ , whereas the variance based methods detect the true scaling only in the Gaussian case. Thus, in the Lévy case, only the DEA reveals the true scaling. The accuracy of all the numerical points is estimated to be of about 1%, for all the cases but the short-time regime of the highly subdiffusional regime studied in Fig.1 and Fig. 2. In this case the error is estimated to be slightly larger, of about 2%. To make easier for the reader to interpret the figures, we have not reported the error bars.

A. Gauss statistics: Fractional Brownian diffusion

Fractional Brownian diffusion is produced by Fractional Brownian noise. We generate a time series $\{\xi_{H,t}\}$ of N data by using the original algorithm by Mandelbrot, illustrated in the book of Feder [21]. Our choice is motivated only by historical reasons, and it is beyond the purpose of this paper to discuss the computational relevance of other algorithms, more recently proposed to generate FBM [8,22]. Chosen a value of $H \in [0 : 1]$, let $\{\theta_i\}$ be a set of Gaussian random variables with unit variance and zero mean. The discrete fractional Brownian increment is given by

$$\xi_{H,t} = x_H(t) - x_H(t-1) = \frac{n^{-H}}{\Gamma(H+0.5)} \left\{ \sum_{i=1}^n i^{H-0.5} \theta_{1+n(M+t)-i} + \right.$$

$$\sum_{i=1}^{n(M-1)} \left((n+i)^{H-0.5} - i^{H-0.5} \right) \theta_{1+n(M-1+t)-i} \Bigg\},$$

where M is an integer that should be moderately large, and n indicates the number of the fractional steps for each unit time. Note that the time t is discrete. In the simulation, good results are obtained with $M = 1000$ and $n = 10$. The time series $\{\xi_{H,t}\}$ is then used for generating a diffusion process with the trajectories (38).

According to the theoretical arguments of Section 3, we expect $\delta = H$ in this case. To confirm this expectation by means of the statistical analysis we generate five different sequences, with the following values of H : (1) $H = 0.8$, (2) $H = 0.6$, (3) $H = 0.5$, (4) $H = 0.4$, (5) $H = 0.2$. We analyze these sequences with the SDA (Fig. 1) and with the DEA (Fig. 2). The results of the numerical analysis fully confirm our expectation. Let us see all this in more detail. For illustration convenience, in Fig. 1 we show the normalized standard deviation defined by

$$NSD(l) = \frac{SD(l)}{SD(1)}. \quad (57)$$

where $SD(l)$ is defined by Eq.(41). With this choice, at $l = 1$ all the numerical results yield, in the ordinate axis, the same value, equal to the unity. For the same reason, in Fig. 2, we plot the entropy difference $S(l) - S(1)$, thereby making all five numerical curves depart from the same ordinate value at $l = 1$. In both figures the straight lines are the results of a fitting procedure, based on $f_{SD}(l) = l^H + K_{SD}$ in Fig. 1, and on $f_E(l) = \delta \ln(l) + K_E$, in Fig. 2. Note that only the levels of the straight lines of these two figures, namely the values K_{SD} 's and K_E 's, whose actual value is of no interest, are determined by the fitting procedure. The parameters H of the straight lines of Fig. 1 are given by the earlier mentioned set of values, selected to create the five artificial sequences. The parameters δ of the straight lines of Fig. 2 are the corresponding scaling values, prescribed by the theoretical remarks of Section 3, namely $\delta = H$ for each curve. The extremely good agreement between theory and numerical experiment, shown in Fig. 1 and Fig. 2, proves that, as expected, in the case of FBM, the condition $H = \delta$ really holds true.

It is remarkable that for almost all the values of H the parameters K_{SD} and K_E are very close to zero. This is a consequence of a peculiar property of FBM. In general, the statistical analysis of times series is affected by the transition from dynamics to thermodynamics. The short-time regime is a kind of dynamic regime and the scaling regime is a kind of thermodynamic regime. It takes time to make a transition from the dynamic to the thermodynamic regime. This property is made especially evident by the analysis of Lévy walk [15]. In the ideal case of FBM, however, the transition time is expected to be equal to zero. This means that, in the case of ideal FBM,

the parameters K_{SD} and K_E should vanish. This constraint is satisfactorily fulfilled by all the curves of Fig. 1 and Fig. 2, but those corresponding to $H = \delta = 0.2$. This seems to be a consequence of the numerical inaccuracy, enhanced by the highly antipersistent condition of this case.

B. Lévy statistics: flight and walk diffusion

We generate five time series of numbers, $\{r_i\}$, distributed according to the following inverse power law

$$\psi(r) = (\mu - 1) \frac{T^{\mu-1}}{(T + r)^\mu}. \quad (58)$$

We select the following values of μ : (1) $\mu = 2.8$, (2) $\mu = 2.6$, (3) $\mu = 2.5$, (4) $\mu = 2.4$ and (5) $\mu = 2.2$. We also select the numbers s_i , equal to 1 or to -1 , randomly according to the coin tossing. We generate the Lévy flight of Section 3B, by building up the new sequence $\{\xi_i\}$, where $\xi_i \equiv s_i r_i$. As to the Lévy walk, we derive it, as explained in Section 3C, by sewing together into a single sequences many patches. These patches are filled with either an integer number of 1's or an integer number of -1 , according to the coin tossing prescription. The distribution density of the patch length is the same as that of Eq. (58). As a consequence, the Lévy flight and the Lévy walk, in the asymptotic time limit have the same scaling, given by Eq.(24). However, as explained in Section 3C, the Lévy walk is expected to result in a given H , predicted by Eq.(34). In Table I we have reported for the reader conveniences the values of δ and H associated to each μ of the five artificial sequences that we are considering for our numerical check.

Fig. 3 shows the DEA at work on the time series generating Lévy flight. The straight lines are fitting functions of the form $f_E(l) = \delta \ln(l) + K_E$. The values of the parameters δ , are not fitting parameters, but are determined by the theoretical prescription of Table I. Figs. 4 and 5 illustrate the results of the SDA and RRA, respectively, applied to the same five time series of Fig. 3. Both figures yield for H a value independent of μ . This value is $H = 0.5$ in all cases. According, to the traditional wisdom, this would suggest the wrong conclusion that we are in the presence of ordinary Brownian motion. We are not, and the DEA is warning us from making this wrong conclusion. The reason for this misleading result is that these techniques are determined by both the finite value of the variance, due to statistical limitation and the memoryless nature of the sequence $\{r_i\}$. The smaller the parameter μ , the smaller the variance, as shown by Fig. 4. The RRA eliminates this spreading, due to the fact that it normalizes the data by dividing by the standard deviation.

Figs. 6-10 refer to the time series generating Lévy walk. Fig. 6 illustrates the result of the DEA. As in

Fig. 3 the straight lines are fitting functions of the form $f_E(l) = \delta \ln(t) + K_E$, and, again the parameters δ are those reported in Table I, and are determined by theoretical prescription $\delta = 1/(\mu - 1)$. Figs. 7, 8, 9 and 10 illustrate the results of RRA, SDA, DFA and SWA, respectively. In these cases the straight lines are fitting functions with the form $f(l) = Kl^H$. The parameters H are not fitting parameters, and their values, reported in Table I, are determined by $H = (4 - \mu)/2$. It is remarkable that, for all these techniques, the same value of μ yields the same value of H , as the fitting curves show. This fully supports the conclusion that these seemingly different methods of analysis are actually equivalent, and that all of them are reliable only in the FBM case. On the other hand, we notice that the values of δ and the values of H reported in Table I fit the condition of Eq.(35) and this a strong evidence that the statistics generated by the time series is Lévy statistics. This means that the disagreement between the DEA and the ordinary techniques of analysis can be used for the important purpose of defining the nature of statistics generated by strange kinetics.

VII. SIGNIFICANCE OF THE RESULTS OBTAINED

This paper affords a compelling evidence that the DEA is the only method leading in all conditions to the detection of the correct scaling exponent δ . In the case of a sequence of random numbers, that according to the GCLT should result in an anomalous scaling, the popular Hurst method would lead to the wrong conclusion that the process observed is equivalent to ordinary Brownian motion. All the traditional methods would lead to quite correct conclusion only in the case of Gaussian statistics, a condition which does not mean, of course, ordinary Brownian diffusion, as made evident by the FBM theory of Mandelbrot. It is also evident that these traditional methods ought not to be abandoned, even if they have to be used with caution. The results of Section 6B prove that the departure of δ from H is a clear indication of the occurrence of Lévy statistics. More in general, the departure of the traditional methods from the DEA finding might be used to shed light on statistics as well as on dynamics. Paraphrasing the title of a recent paper [23], "Do strange kinetics imply unusual thermodynamics?", we can say that one of the basic problems concerning complex systems is that of establishing if anomalous diffusion (strange kinetics) is compatible or not with ordinary Gaussian distribution (ordinary thermodynamics). The connection between thermodynamics and statistical equilibrium does not need any explanation. Here we are subtly assuming that scaling can be perceived as a form of thermodynamic equilibrium. If we look to the results of this paper from within this per-

spective, we can conclude that FBM is an example of strange kinetic compatible with ordinary thermodynamics. We can thus conclude that the joint use of DEA and techniques of analysis based on variance can assess when strange kinetics force the complex system to depart from ordinary thermodynamics.

-
- [1] H.E. Stanley, L.A.N. Amaral, P. Gopikrishnan, P. Ch. Ivanov, T.H. Keitt, V. Plerou, *Physica A* **281** 60 (2000).
 - [2] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Perseus Book, Reading, Massachusetts, 1985).
 - [3] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldenberger, *Phys. Rev. E* **49**, 1685 (1994).
 - [4] A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics* (Dover Publications, Inc. New York, 1949).
 - [5] N. Scafetta, P. Hamilton, P. Grigolini, *Fractals* **9**, 193 (2001).
 - [6] P. Allegrini, M. Barbi, P. Grigolini and B.J. West, *Phys. Rev. E* **52**, 5281 (1995).
 - [7] H. E. Hurst, R. P. Black, Y. M. Simaika, *LongTerm Storage: An Experimental Study* (Constable, London, 1965).
 - [8] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge (2000).
 - [9] B.B. Mandelbrot, *The Fractal Geometry of Nature*, (Freeman, New York, 1983).
 - [10] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II* (Springer-Verlag, Berlin, 1991).
 - [11] M. Annunziato, P. Grigolini and J. Riccardi, *Phys. Rev. E* **61**, 4801 (2000).
 - [12] B.J. West, P. Grigolini, R. Metzler, T.F. Nonnenmacher, *Phys. Rev. E* **55** (1997) 99.
 - [13] S. Jespersen, R. Metzler, H. C. Fogedby, *Phys. Rev. E* **59** (1999) 2736.
 - [14] B.V. Gnedenko, A.N. Kolmogorov, *Limit Distributions for Sum of Independent Random Variables*, (Addison Wesley, Reading 1954).
 - [15] P. Grigolini, L. Palatella, G. Raffaelli, *Fractals*, **9**, 439 (2001).
 - [16] T. Geisel, J. Nierwetberg, and A. Zacherl, *Phys. Rev. Lett.* **54**, 616 (1985).
 - [17] M.F. Shlesinger, J. Klafter, *Phys. Rev. Lett.* **54**, 2551 (1985).
 - [18] M. F. Shlesinger, B.J. West, J. Klafter, *Phys. Rev. Lett.* **58**, 1100 (1987).
 - [19] T. Geisel and S. Thomae, *Phys. Rev. Lett.* **52**, 1936 (1984).
 - [20] K.Hu, P. Ch. Ivanov, Z. Chen, P. Carpena, H.E. Stanley, *Phys. Rev. E* **64**, 011114 (2001).
 - [21] J. Feders, *Fractals*, Plenum Publishers, New York (1988).
 - [22] Hernán A. Makse, Shlomo Havlin, Moshe Schwartz, and H. Eugene Stanley, *Phys. Rev. E* **53**, 5445 (1996).
 - [23] I.M. Sokolov, J. Klafter, and A. Blumen, *Phys. Rev. E*

64 021107 (2001).

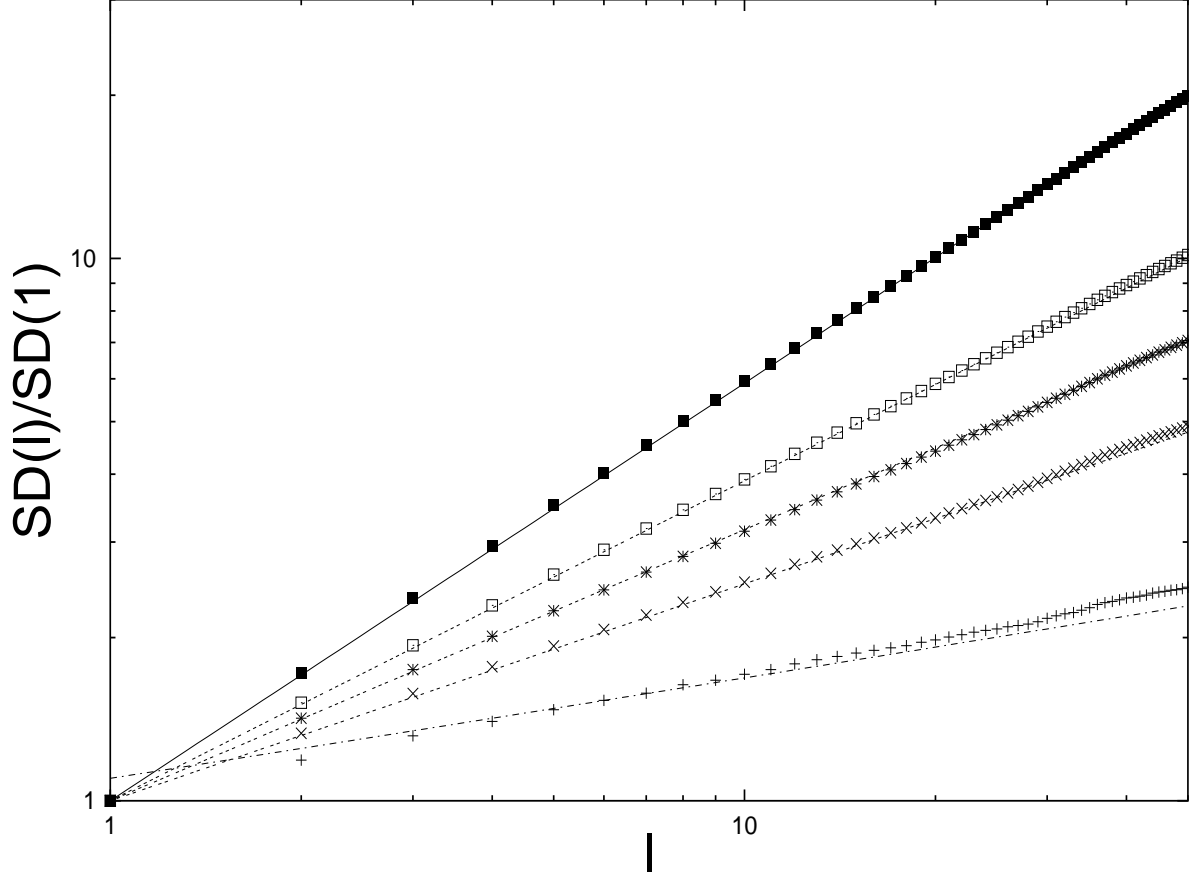


FIG. 1. SDA in action on the five time series of Fractional Brownian noise of Section 6A. In ordinate we plot $SD(l)/SD(1)$ as a function of l . The straight lines are fitting functions with the form $f_{SD}(l) = l^H + K_{SD}$. The values of H are not the result of the fitting procedure, but they correspond to the set of values used to create the five artificial sequences of Section 6 A. From the top to the bottom these values are: (1) $H = 0.8$; (2) $H = 0.6$; (3) $H = 0.5$; (4) $H = 0.4$; (5) $H = 0.2$.

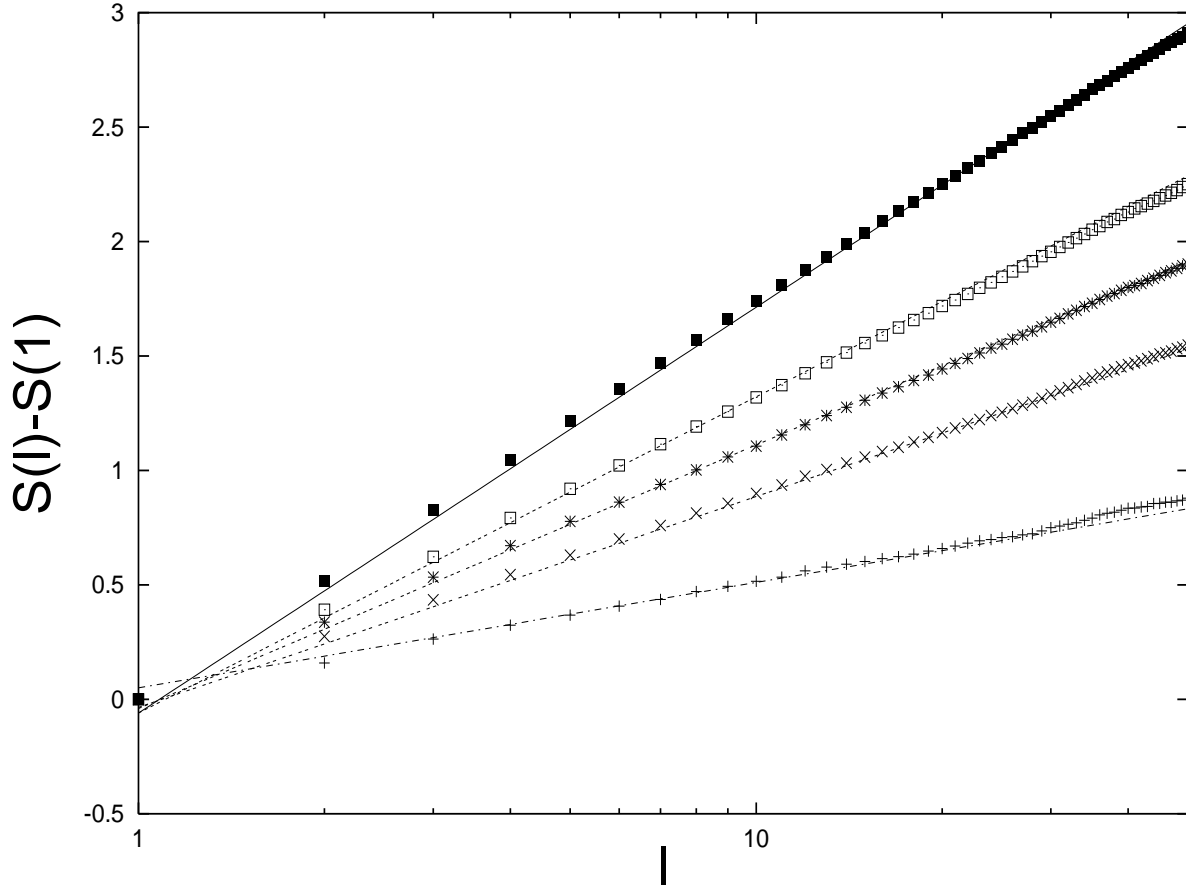


FIG. 2. DEA of the five time series of Fractional Brownian noise of Section 6A. For illustration convenience, in ordinate we plot, as a function of l , the entropy increment $S(l) - S(1)$. The straight lines are fitting functions with the form $f_E(l) = \delta \ln(t) + K_E$. The values of H are not the result of the fitting procedure, but they correspond to the set of values used to create the five artificial sequences of Section 6 A, according to the theoretical prescription $\delta = H$. From the top to the bottom these values are: (1) $\delta = 0.8$; (2) $\delta = 0.6$; (3) $\delta = 0.5$; (4) $\delta = 0.4$; (5) $\delta = 0.2$.

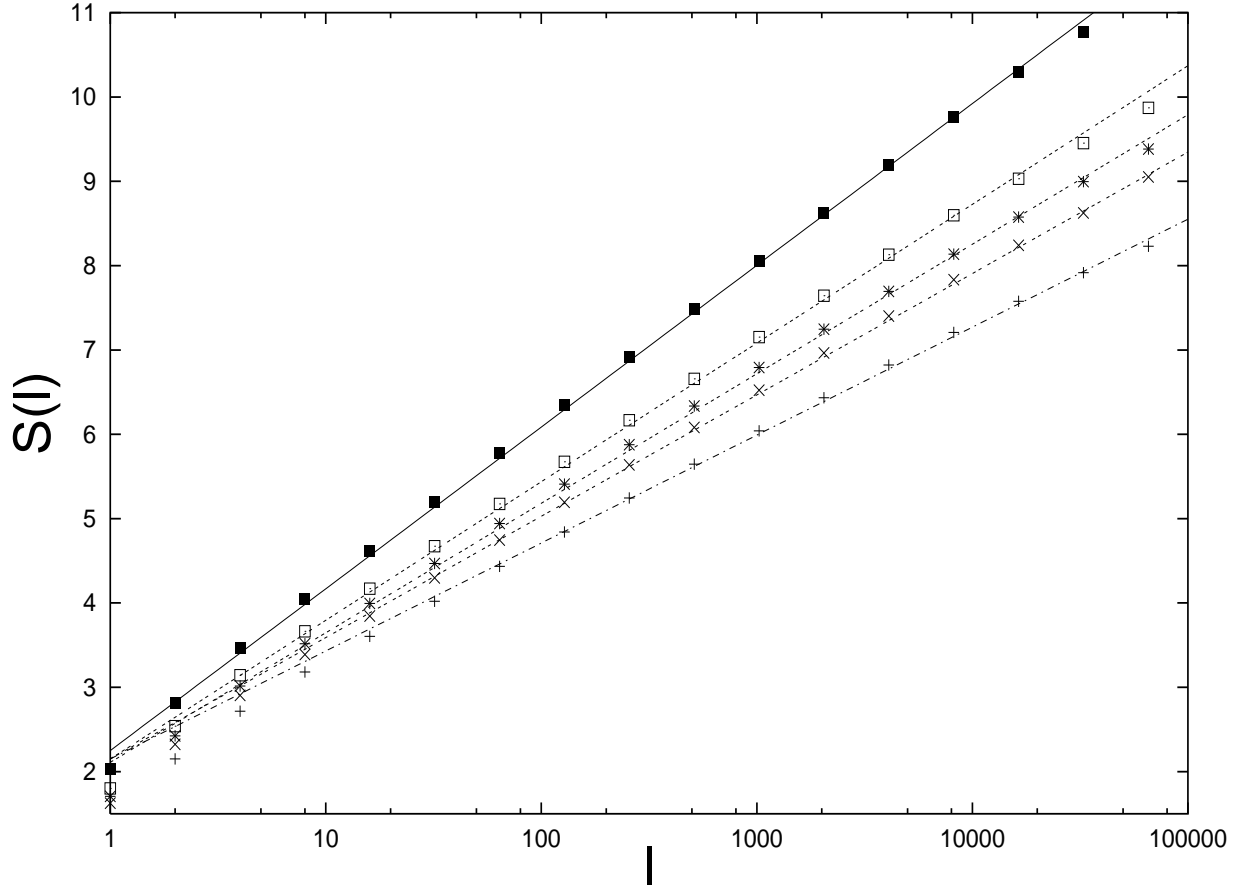


FIG. 3. DEA of the five Lévy flight time series of Section 6B. The straight lines are fitting functions with the form $f_E(l) = \delta \ln(l) + K_E$. The values of δ are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6 B, according to the theoretical prescription $\delta = 1/(\mu - 1)$. From the top to the bottom these values, and the corresponding fitting parameters K_E , are: (1) $\delta = 0.833$, $K_E = 2.25$; (2) $\delta = 0.714$, $K_E = 2.15$; (3) $\delta = 0.667$, $K_E = 2.11$; (4) $\delta = 0.625$, $K_E = 2.15$; (5) $\delta = 0.556$, $K_E = 2.15$.

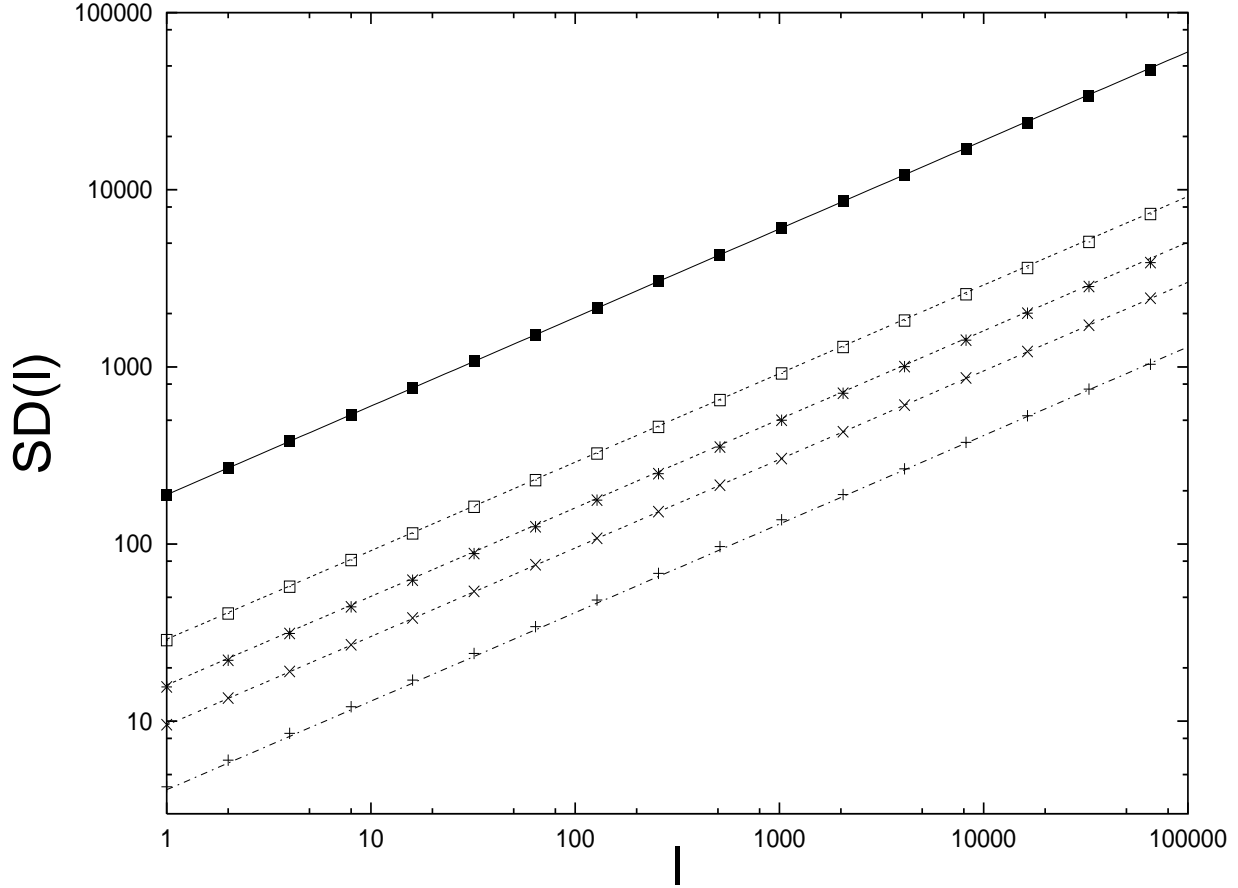


FIG. 4. SDA of the five Lévy flight time series of Section 6B. The straight lines are fitting functions with the form $f_{SD}(l) = K_D l^H$. From the top to the bottom we have: (1) $H = 0.5, K_D = 190$; (2) $H = 0.5, K_D = 29$; (3) $H = 0.5, K_D = 16$; (4) $H = 0.5, K_D = 9.5$; (5) $H = 0.5, K_D = 4.1$.

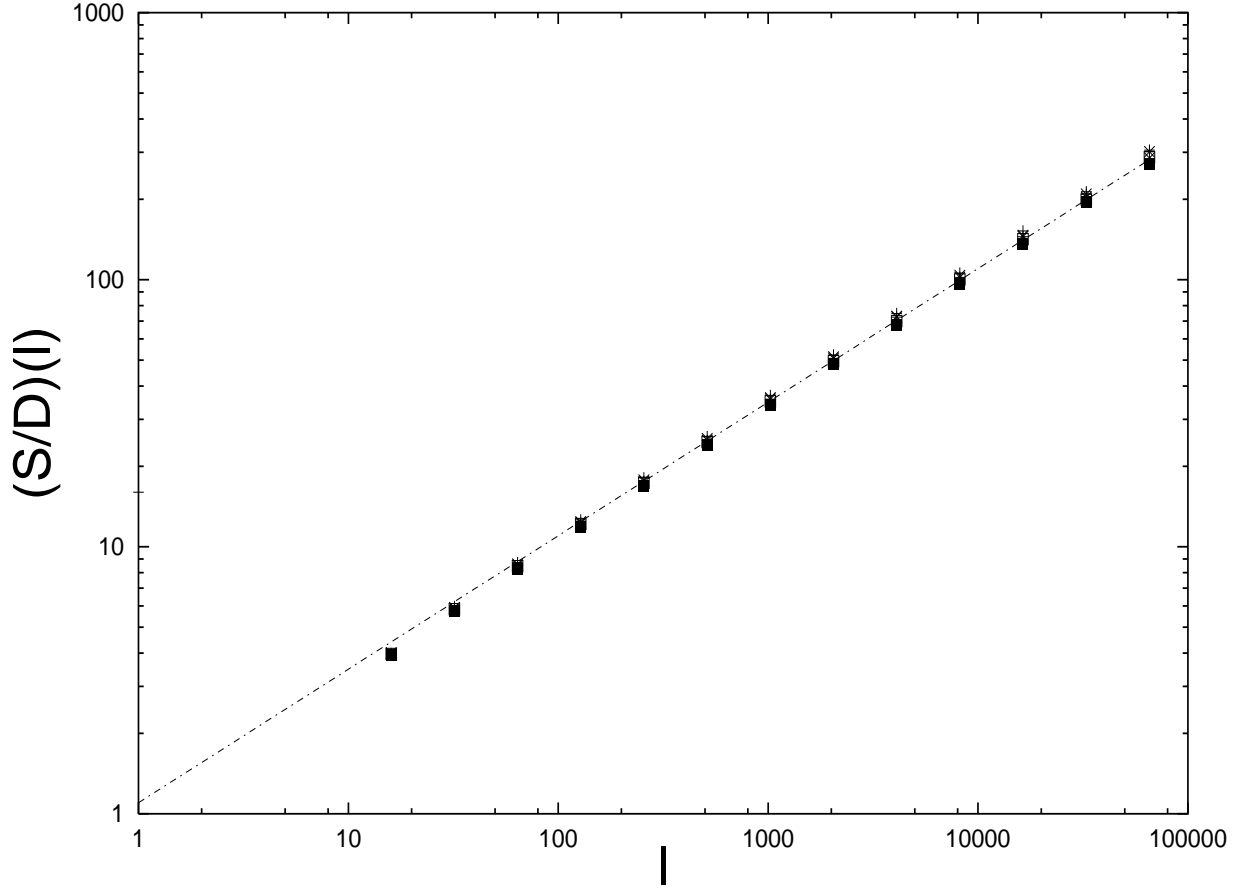


FIG. 5. RRA of the five Lévy flight time series of Section 6B. All the five cases are fitted by the straight line corresponding to the fitting function $f_{S/D}(l) = K_{S/D} l^H$, with $K_{S/D} = 1.1$ and $H = 0.5$.

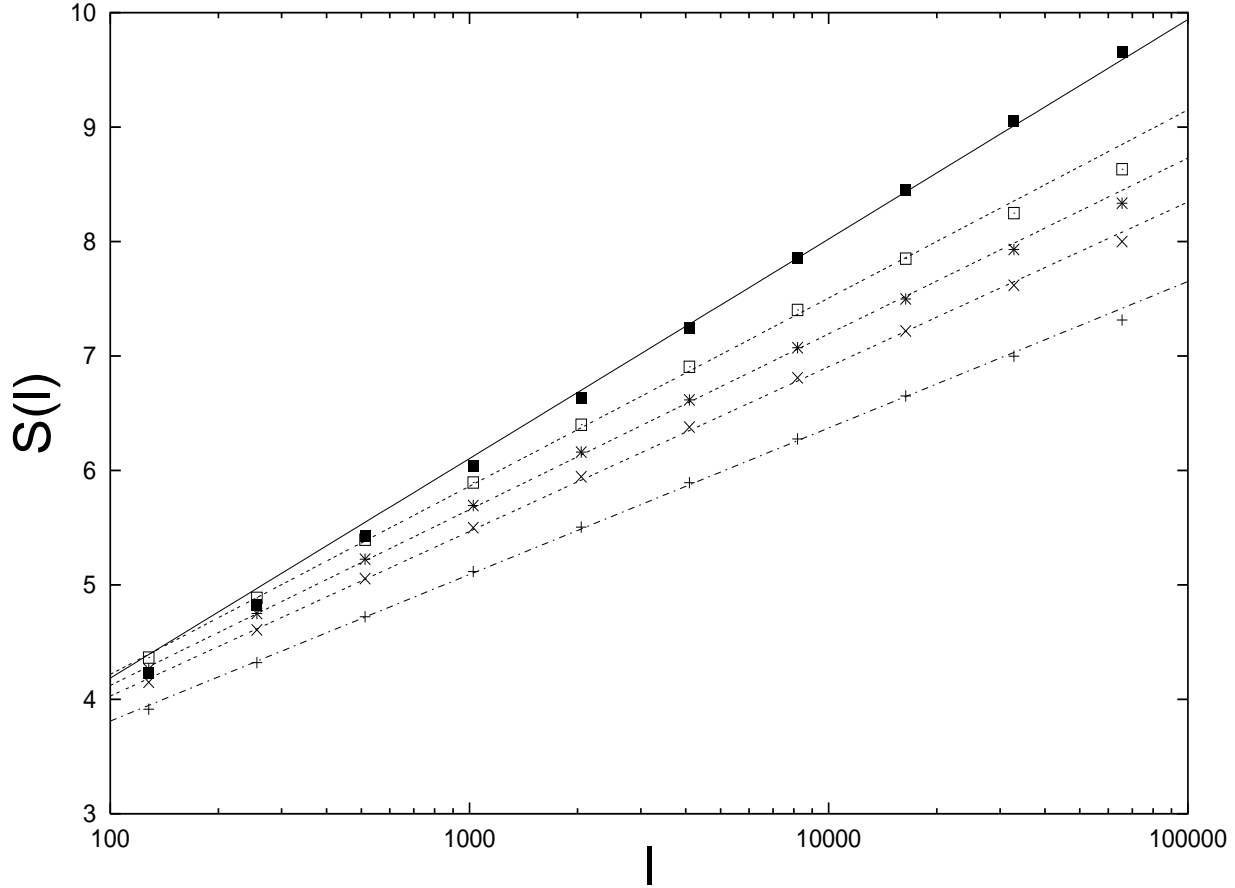


FIG. 6. DEA of the five Lévy walk time series of Section 6B. The straight lines are fitting functions with the form $f_E(l) = \delta \ln(l) + K_E$. The values of δ are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6B, according to the theoretical prescription $\delta = 1/(\mu - 1)$. From the top to the bottom these values, and the corresponding fitting parameters K_E , are: (1) $\delta = 0.833$, $K_E = 0.35$; (2) $\delta = 0.714$, $K_E = 0.93$; (3) $\delta = 0.667$, $K_E = 1.05$; (4) $\delta = 0.625$, $K_E = 1.15$; (5) $\delta = 0.556$, $K_E = 1.25$.

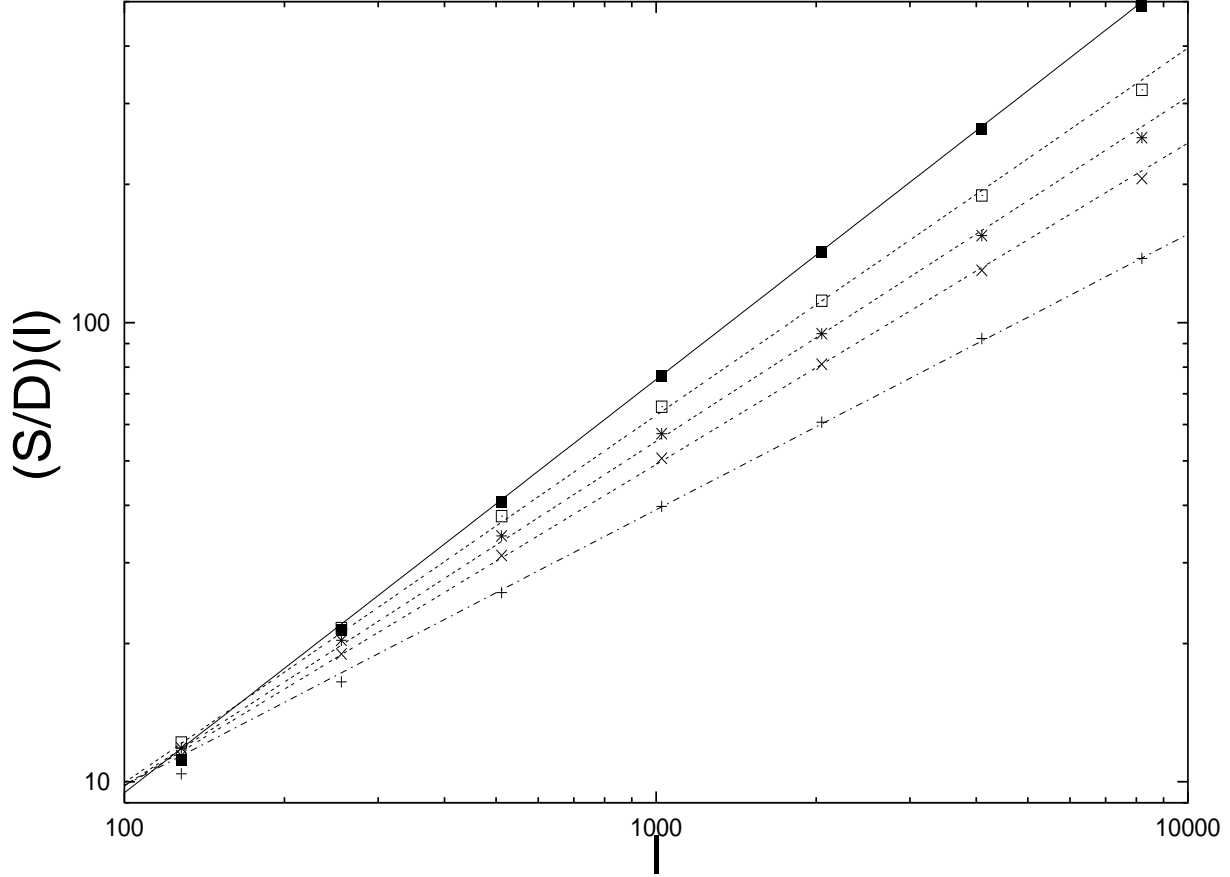


FIG. 7. RRA of the five Lévy walk time series of Section 6B. The straight lines are fitting functions with the form $f_{S/D}(l) = K_{S/D} l^H$. The values of H are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6B, according to the theoretical prescription $H = (4 - \mu)/2$. From the top to the bottom these values, and the corresponding fitting parameters $K_{S/D}$, are: (1) $H = 0.9, K_{S/D} = 0.15$; (2) $H = 0.8, K_{S/D} = 0.25$; (3) $H = 0.75, K_{S/D} = 0.75$; (4) $H = 0.7, K_{S/D} = 0.39$; (5) $H = 0.6, K_{S/D} = 0.62$.

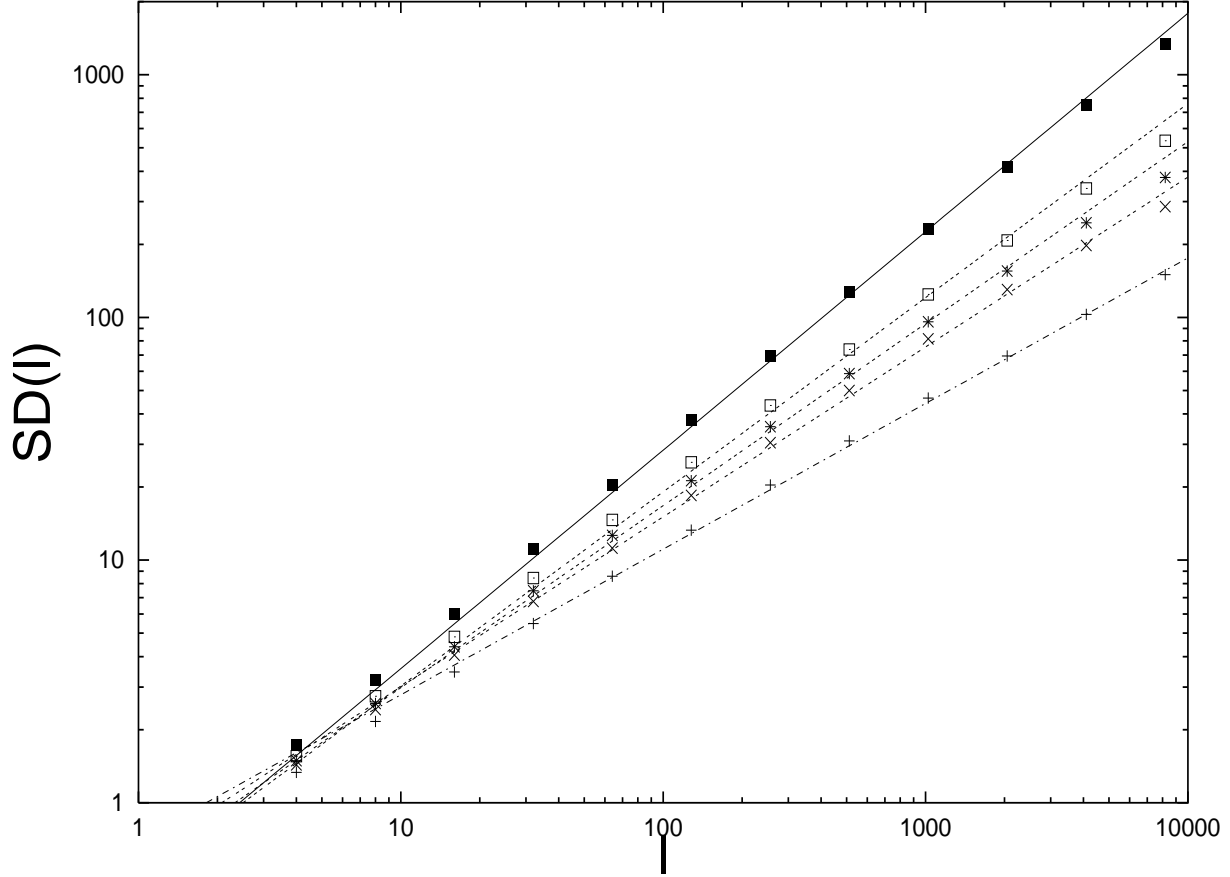


FIG. 8. SDA of the five Lévy walk time series of Section 6B. The straight lines are fitting functions with the form $f_{SD}(l) = K_{SD} l^H$. The values of H are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6B, according to the theoretical prescription $H = (4 - \mu)/2$. From the top to the bottom these values, and the corresponding fitting parameters K_D , are (1) $H = 0.9, K_D = 0.45$; (2) $H = 0.8, K_D = 0.48$; (3) $H = 0.75, K_D = 0.53$; (4) $H = 0.7, K_D = 0.6$; (5) $H = 0.6, K_D = 0.7$.

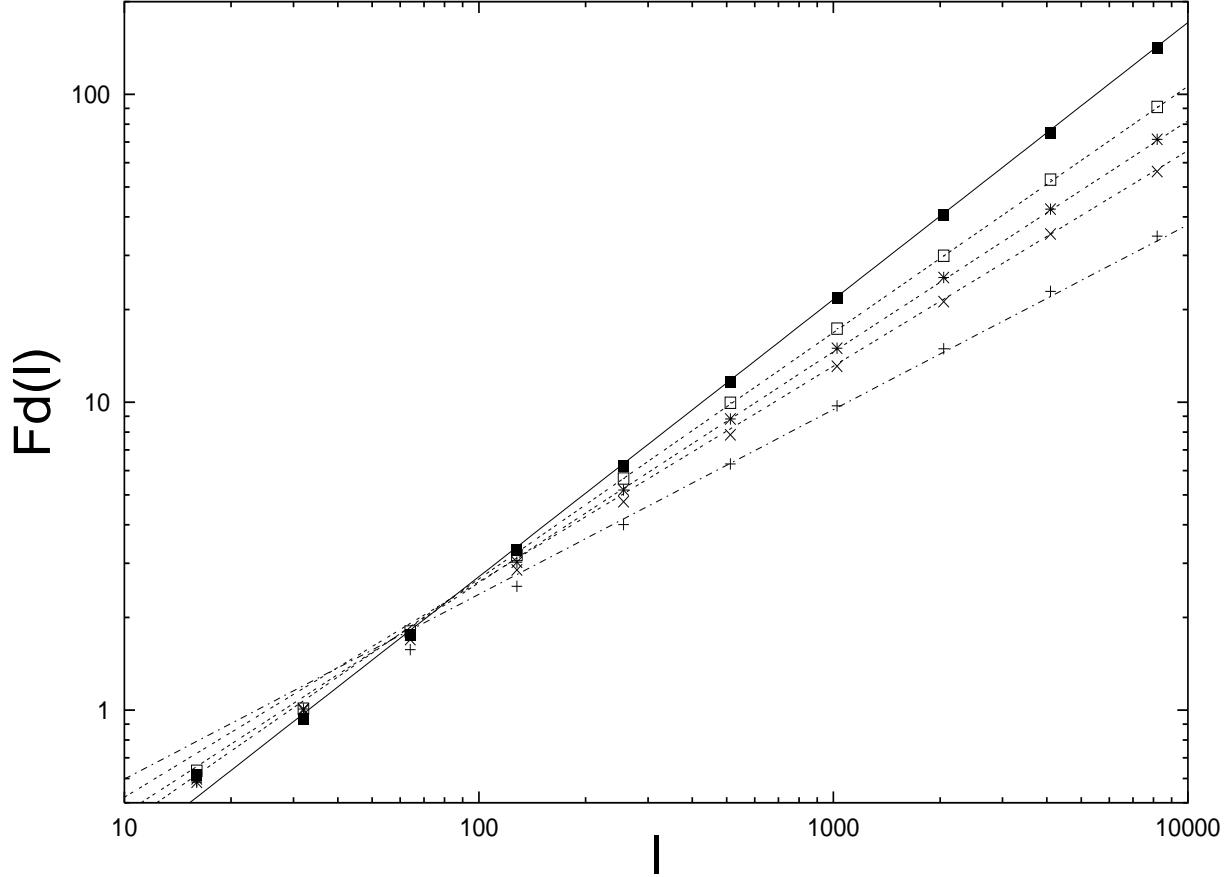


FIG. 9. DFA of the five Lévy walk time series of Section 6B. The straight lines are fitting functions with the form $f_F(l) = K_F l^H$. The values of H are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6B, according to the theoretical prescription $H = (4 - \mu)/2$. From the top to the bottom these values, and the corresponding fitting parameters K_F , are: (1) $H = 0.9, K_F = 0.043$; (2) $H = 0.8, K_F = 0.067$; (3) $H = 0.75, K_F = 0.082$; (4) $H = 0.7, K_F = 0.104$; (5) $H = 0.6, K_F = 0.15$.

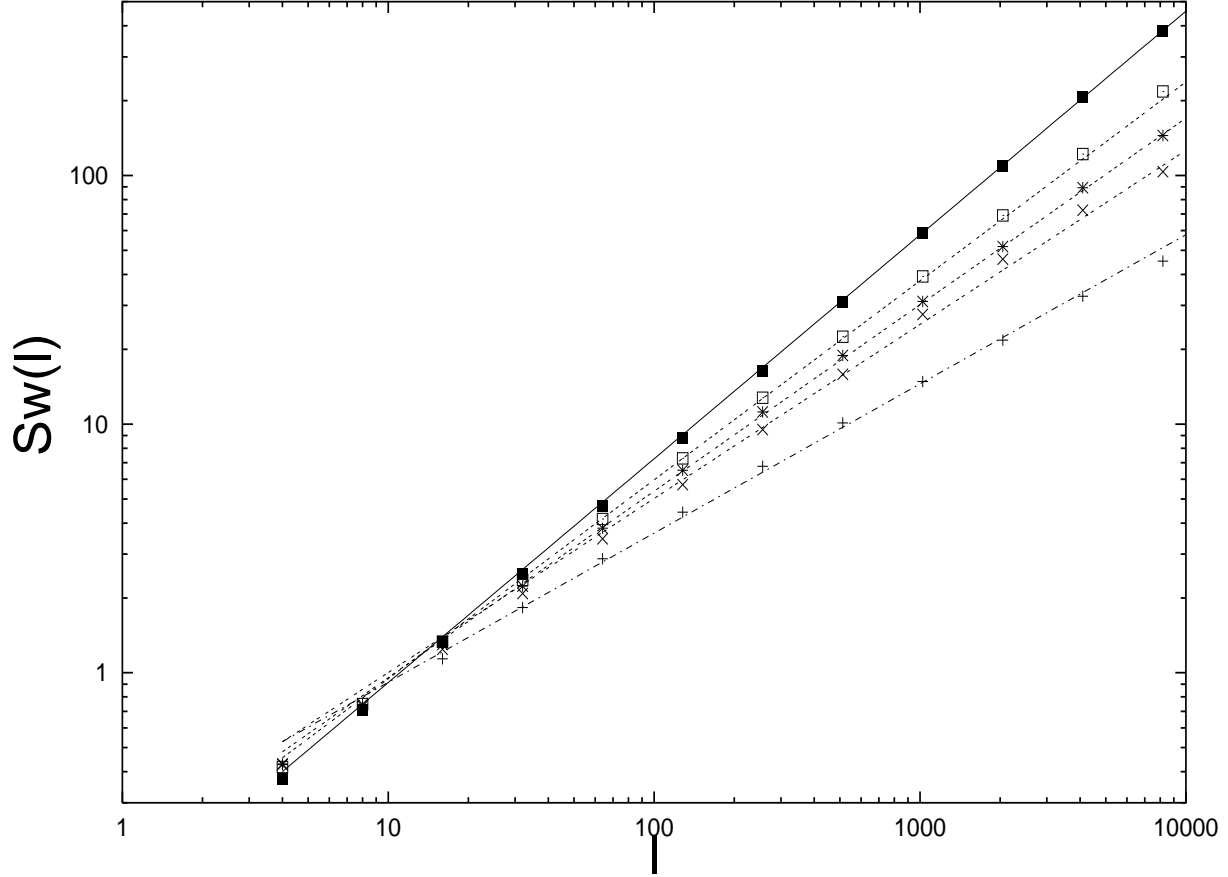


FIG. 10. SWA of the five Lévy walk time series of Section 6B. The straight lines are fitting functions with the form $f_W(l) = K_W l^H$. The values of H are not the result of the fitting procedure, but they correspond to the set of values $\{\mu\}$, used to create the five artificial sequences of Section 6B, according to the theoretical prescription $H = (4 - \mu)/2$. From the top to the bottom these values, and the corresponding fitting parameters K_W , are: (1) $H = 0.9, K_W = 0.115$; (2) $H = 0.8, K_W = 0.15$; (3) $H = 0.75, K_W = 0.17$; (4) $H = 0.7, K_W = 0.2$; (5) $H = 0.6, K_W = 0.23$. The wavelet spectral density is calculated using the Maximum Overlap Discrete Wavelet Transform with the Daubechies H4 discrete wavelet [8].

μ	H	δ
2.2	0.90	0.833
2.4	0.80	0.714
2.5	0.75	0.667
2.6	0.70	0.625
2.8	0.60	0.556

TABLE I. Theoretical relation between the waiting time distribution power exponent μ and the variance scaling exponent H and the pdf scaling exponent δ .